P9102

William J. Craig
Center for Urban & Regional Affairs
University of Minnesota
Minneapolis, MN 55455

EMPTY PROMISES FOR SMALL AREA DATA: MONITORING OUR NEIGHBORHOODS USING OPERATIONAL RECORDS[1]

Abstract: In the late 1960s, the federal USAC project, coupled with new Census Bureau address-matching technology from the New Haven project, raised the hope that operational records could be used by cities to monitor changes within their neighborhoods. This would allow us to act immediately when problems arise, not having to wait ten years between censuses. Two decades later little more current neighborhood information is available despite enormous increases in computerized data and geocoding technology. Using Minneapolis and St. Paul as a case study, very little current data has been published for sub-city areas. Even where the raw data exists, technical and institutional barriers prevent generation of summary reports for small areas. Bureaucratic inertia has led to the design of computer systems and departmental rules focussed on serving the operational needs of each office separately, without reguard to the potential value of that department's data to other analysts, managers, or decision makers.

INTRODUCTION

The 1990 Census is beginning to be delivered and the drought of small area data is at an end. With a few exceptions, mostly related to housing, subcity data are more than a decade old. City programs and neighborhood initiatives have been flying blind for ten years. They have been unable to monitor points of deterioration or evaluate the effectiveness of interventions. We were promised a better situation, but those promises remain unfulfilled.

Those promises go back over two decades. Colley (1974) argued for the need for a current tract level social index of urban pathology. A paradigm was developing that the daily operational records of government should be summarized and synthesized to provide necessary information for management and decision making. Based on this paradigm the federal government funded the multi-million dollar USAC (Urban Systems Inter-Agency Committee) project to prototype urban information systems in the early 1970s. Researchers argued that "[t]here is impressive replicable evidence to justify the view that operational data can satisfy a large majority of known and established planning data requirements...." (Mitchell 1971, 138). Much of this synthesis was based on summarizing addressed-based records to census tract and other geographies. The technology for matching addresses to tracts, called

---

[1]This paper was presented at the 1991 Applied Geography Conference in October. It is a revised version of one with the same title presented at the 1991 Conference of URISA (the Urban and Regional Information Systems Association) in August.

geocoding, was developed at the Census Use Project in New Haven, to facilitate a mail census --used for the first time in 1970 (Cooke and Maxfield 1967).

In fact, the current picture of the availability of small area population and housing data is not much better than it was 20 years ago, at least if Minneapolis and St. Paul are good examples.[2] Some progress has been made in the Twin Cities, but that progress has not been uniform and some of the gains have been lost subsequently.

This paper looks at small area data in two different ways: 1) existing reports and 2) the capability to generate special data summaries on demand. In the first instance we explore the leadership role of government (and private concerns) in generating information about small areas. In the second, we look at the responsiveness of government agencies (and others) to the needs of the community. This paper keeps a focus on the types of data found in the Census of Population and Housing, and not on the types of data found in the various economic censuses, primarily because the focus is on people and neighborhoods.

EXISTING REPORTS

Existing small area reports (published and unpublished) come from a variety of sources, are produced on many different time schedules, and subdivide cities in a variety of ways. The University of Minnesota was approached by a local foundation in the spring of 1990 to help identify marginal neighborhoods that might benefit from a set of comprehensive services. This section documents the results of attempting to find current data to characterize neighborhoods.

One of the problems facing researchers looking for small area data is the variety of ways to divide a city. Table 1 shows six existing ways of publishing data in Minneapolis and St. Paul, the number of such areas in each city, and the average size of each area. Census tracts have many advantages, but their boundaries do not correspond with citizen participation areas. Minneapolis and St. Paul use vastly different size areas for citizen participation, though Minneapolis does authorize some functions at the comparable district ("community") level. St. Paul does not have small neighborhood areas, but substitutes data for quarter sections; these areas are based on the Public Land Survey and are square areas one-half-mile on a side. A few ZIP areas extend past the municipal boundaries. The Multiple Listing Service (MLS) has data on the real estate market. With rare exceptions, these various areas are not coterminous; *i.e.* their boundaries do not match.

TABLE 1: TYPE AND SIZE OF ALTERNATIVE SUB-CITY AREAS
(Recognized Citizen Participation Areas Shown in **Bold**)

	Minneapolis		St. Paul	
	# Units	Avg. Size	# Units	Avg. Size
Total population	1	368,383	1	272,235
Census tracts	126	2,924	82	3,320
Neighborhoods	**85**	**4,334**	-	-
Districts	11	33,489	**17**	**16,014**
Quarter sections	-	-	205	1,328
ZIP areas	22	16,745	12	22,686
MLS districts	10	36,836	12	22,686

[2] St. Paul was one of the six cities across the county participating in USAC. St. Paul efforts focussed on human services, an area closely related to the population and housing data of this study.

The results of the search for printed and unprinted reports are shown in Table 2[3]. With the exception of tract data on the length of residence, neither city has any current population data below the city level. Both have abundant housing data, which are good for monitoring neighborhood change, but St. Paul's data come sporadically and are less available at scales which might aid neighborhood groups. Some additional rent data are available from a private Apartment Guide service. Income data for residents are nonexistent, jobs and welfare data are relatively good, but unemployment figures are really only good at the city level. Data about crime in neighborhoods are good. Some of these data were collected at for one type area, but reported at another, not always with adequate documentation.

It would be a reasonable summary to say that in the better case, Minneapolis, current neighborhood data are reported for dwellings, welfare, crime, and problem births and pregnancies. In the lesser case, St. Paul, most of those same types of information are also available, but less frequently, and at scales which do not match citizen participation districts. Neither has adequate population, income, or unemployment data. These three topics are relatively critical measures of neighborhood stability. These topics will be missing in most cities, for example, they are absent from Chicago's Neighborhood Information System Project (Greer 1991). The reasons stem from 1) the fact that local government's major responsibility is to handle real property data and 2) lack of interest from higher levels of government in city or subcity data. Because of the lack of current data for the two cities, our report to the foundation ultimately was derived from the 1980 census -- ten year old data.

ACCESS TO OPERATIONAL RECORDS

Early leaders thought that data from day-to-day operations could be summarized to provide effective data for management and policy making. The experiences in Minneapolis and St. Paul show that there are still many barriers to overcome before this dream can become a reality.

---

[3]Sources of this information are given below –

Minneapolis:	*State of the City,* annual
	Comprehensive Law Enforcement Plan, 1989
	Real Estate Activity Report, annual
	Neighborhood Profile, Whittier Neighborhood, 1990
St. Paul:	*St. Paul Today,* 1986
	District Profiles, 1985
	Housing Policy for the 1990s, 1990
	Crime Analysis Report, annual and 1985-1990 summary
State and Metropolitan:	*Population Household Estimates,* annual
	Twin Cities Area Average Covered Employment, quarterly
	PK-12 Student Ethnic Enrollment, annual
	Housing Vacancies and Turnovers in the Twin City Metropolitan Area, quarterly
	Local Area Unemployment Statistics, monthly

TABLE 2: DATA AVAILABILITY FOR INTERCENSUS PERIODS

Table shows the smallest units available for each city from published reports and a code for the frequency of reporting: 1=annually or more frequently
2=2 to 5 years
3=sporadically

	Mpls	St. Paul	Comment
Population			
Total count	city-1	city-1	estimate by regional government
Race/ethnicity	city-1	city-1	estimates from school district data
Age	none	none	
Mobility	tract-1	city-1	based on electric utility connections
Housing			
Number of units	neighbor-hood-1	city-2	updates based on building permits
Tenure	district-1	qtr sec-3	Based on tax homestead status. St. Paul reported for 1987 only.
Rent	city-1	city-2	based on newspaper ads and a commercial guide, therefore biased towards newer and higher cost units
Apt vacancy rate	tract-1	city-1	based on electric utility data;
Value	district-1	qtr sec-3	Based on sales Certificate of Value. St. Paul reported for 1988 only. MLS data is available annually.
Size of units	none	none	
Condition	district-1	district/tract-3	Mpls uses assessor ratings St. Paul surveyed in 1981 and 1988.
Income/Employment			
Income	none	none	Minneapolis had tract level data through 1984.
Jobs	ZIP/districts-1	ZIP/districts-1	Some grouping of ZIPs and districts occurs
Unemployment rate	city-1	city-1	Mpls attempts to allocate city data to districts
Welfare	ZIP-1	district-2	Mpls had tract data until 1987; St. Paul district reports are based on tract level data
Crime	neighbor-hood-3	city-1; district/ qtr sec-2	Minneapolis had annual neighborhood level data until 1985/86.
Health	district-2	city-2	The state produces tract level data annually on problem pregnancies. These small numbers are aggregated to permit comparisons.

Cases

Below are listed eight examples of attempts to extract and summarize small area data from existing databases. Each failed for the reasons listed. Some of the reasons are technical, others reflect institutional problems.

The first two examples given below come from a recent research effort to help a public interest group (MICAH--the Metropolitan Interfaith Council on Affordable Housing) create a housing profile of one inner-city neighborhood. A graduate student was hired and other resources made available to dig out data. She had more time and persistence than most people, yet often came up empty. The six other cases come from a variety of other research projects and discussions.

- In the MICAH project, CURA needed to prepare a housing profile of one neighborhood. An important issue for family housing is the number of bedrooms in multi-family units. This information exists in the assessor's data file, but the assessor was unwilling to authorize computer runs, even for a fee. They did allow us access to a computer terminal, which allowed us to find the information (one parcel at a time), but access required either an owner's name or a parcel number. The Minneapolis Planning Department cooperated by listing the parcel numbers of all 697 multi-family buildings. It took the research assistant four days at the terminal to compile the count of bedrooms for the neighborhood. Many bureaucracies limit access to protect confidentiality. Here, because the bureaucrats did not want to do something beyond their normal jobs, an outside person was encouraged to browse through the bedrooms of nearly 700 households.

- On this same project, researchers wanted to know what proportion of recent sales had been facilitated by bank financing, as opposed to seller-financed contracts-for-deed. County computer records had this information only for the current year and access was again limited to one property at a time. Printed reports are available, but are organized by subdivision; the Whittier neighborhood covers all or parts of 59 subdivisions. It took another four working days to go through those records and extract the required data from the abstract office. Even then the nature of the data was cloudy, many identified with a code "county mortgage," in addition to the expected contract and mortgage. No consistent definition could be gained about what this new code might mean. County officials suggested that a robust classification could be made by looking up the transactions on microfilm; during the 1980-87 period there were 2,615 transactions in Whittier and we declined. In Hennepin County, roughly half the parcels are recorded under the abstract system and half under the Torrens system; access to the Torrens records is more convoluted and deemed not worth the effort. In this case many things have gone wrong. Old records are not kept for trend analysis. No mechanism could be found to purchase summary processing. Data items that were not central to the mission of the organization were treated in a sloppy manner.

- Until recently, Minneapolis was able to obtain median family income estimates for its tracts from the Minnesota Department of Revenue for a fee (Heath, Graham, and Nelson 1986). There was always a two-year lag. The department insisted on doing its own computer processing and address-matching, primarily to preserve confidentiality. Since 1986, these data have not been available. Due to simplification of the state income tax forms, data used to estimate family income are no longer in the state computer. Federal taxable income has replaced adjusted gross income as the starting point for computing state taxable income. More importantly, dependent credits have become muddled with old age and disability credits so it is impossible to identify family records with certainty. Filers are

required to enclose copies of their federal forms which have the requisite data, but as a cost saving measure, these data are no longer computerized.

- The Minnesota Department of Jobs and Training uses the required unemployment insurance forms from firms to estimate the number of jobs in the state and reports these data by areas as small as ZIP code area. This is a useful piece of information, flawed only by the fact that some companies report all employees from a headquarters location regardless of the job location. As part of a large traffic analysis study, the Metropolitan Council requested this data by traffic assignment zones (TAZ). The charge to meet this request is $200,000. Much of the cost is due to the database needing to be cleaned up. Still, the annual fee for such data, after the cleaning, is estimated to be $60,000. Since the firms do not move frequently and the TAZ are stable, one must assume that the high fees have been set to gouge a user and to discourage similar requests.

- Schools in both Minneapolis and St. Paul have address and racial information on each pupil so they can develop effective desegregation programs through bussing. Neither system is anxious to share this data, even summarized at a sub-city level, such as tract. A year before the 1990 Census was released, the University of Minnesota began a research project dealing with the issues of knowledge, attitudes, and behavior in the Black community with respect to AIDS. The University needed to know the pattern of post-1980 expansion of the Black community. Minneapolis produced a generalized location map which was used as a guideline for the field interviewers. St. Paul promised several times to get back to us but never did. A solid "no" would have been appreciated.

- Ramsey County, responsible for ownership and assessor records in St. Paul, has no way of determining the census tract of any property in its jurisdiction. Property records have a field for census tract, but it has not been used for decades. Many properties have data in this field, but the numbers have no relationship to tract numbers used recently in St. Paul; they appear to be 1950 tract numbers. As of 1990, there was no mechanism in government to address-match these properties to tract. This explains why so much of St. Paul's published data are available only for the city as a whole. Only because the parcel number is based on the Public Land Survey System are some data available for smaller areas, in this case the quarter-section.

- Minneapolis has excellent opinion data from homeowners. A "homesteaded" property can be taxed at one-third the rate of a rental property, so homeowners are keen to respond. Each year's homestead application comes with a short survey about some topic and data are released at the neighborhood level. The same cannot be done for renters. There is no list of apartment addresses available in either city, so mail surveys are impossible. The building address and address of the owner are important for managing government property operations, but there is no use for apartment numbers, so no records are kept.

- The Regional Multiple Listing Service maintains a database of all properties listed in the metropolitan area. A small but useful portion of this data (*e.g.* average prices and method of financing) is published by Minneapolis Area Association of Realtors, and the St. Paul association is more than happy to provide data summaries. St. Paul even uses subdistricts--which would increase the detail of data available three-fold, but the address-matching algorithm wrongly assigns subdistrict number "1" to any property where the realtor does not specify the location. University researchers have approached the regional board about generating summary tables for different data items and different geographic areas.

Some of the fields of interest to researchers and neighborhood groups have not been reported thoroughly (*e.g.* school district number and finished living area), but generally this database still could provide a wealth of useful information. The regional board liked the idea of providing a community service, but the request is in limbo for two reasons: concern over confidentiality, which should not be an issue for summary reports, and technical/contractual limitations. The regional board contracts out its data processing and the provider limits use by: 1) not being able to address-match records to census tracts, and 2) limiting use of the data to those specified in the original contract. In a sense, the data processing service has become the owner of the data and is taking a narrow view about its use. While these issues are debated, data are being lost. The database is for active listings and recent sales only and no archival records are maintained.

Summary of Barriers to Use

The eight situations described above share three common problems. These are problems with technology, unreasonable concerns with confidentiality, and bureaucratic inertia.

The two major technical problems are inadequate retrieval systems and lack of capacity to geocode addresses. Several systems were designed to meet operational needs and therefore provide access to only one record at a time. The lack of capacity to geocode records, then summarize those records by small area, feeds directly into the second problem.

Protection of confidentiality was sometimes given as a reason to restrict access, even though the request was for summary data, not individual records. The Census Bureau has developed excellent procedures to ensure everyone's confidentiality while providing a wide range of useful summary data. Neighborhood level data are no different, but because bureaucrats do not understand small area data, they attempt to restrict access. This leads to the third concern.

Bureaucratic inertia means people focussing on the mission of their organization where a minor amount of additional work would yield great benefits for others. Several examples are seen where data items have been neglected or historical files discarded because they are not of central value to the organization. Other examples are seen where barriers have been erected to prevent outsiders from using agency resources. Lack of resources to respond was sometimes given as a reason, but refused offers of payment and excessively high charges lead one to believe that bureaucratic initertia is the real reason.

Signs of Hope

The bad experiences that have been encountered in the attempts to access small area data in Minneapolis and St. Paul are only one side of the story. There are hopeful signs as well. This section will list some of those signs. Many of these same breakthroughs are occuring in other cities across the country.

The data provided by the local electric utility have been useful for monitoring apartment vacancies and mobility patterns. Over the years their database has become evermore clean and address matches now have very few rejections.

People are becoming aware of address-matching technology as prices drop and the technology becomes more accessible. For example, a fellow researcher at the University of Minnesota has acquired the MapInfo software package and address

range files for the Twin Cities area for well under $1000; he is using this technology to process data that formerly took him weeks of work.

One of the very few non-census federal datasets available for census tracts has been loan data resulting from the Home Loan Disclosure Act (HMDA). Until this year that dataset was flawed in two ways: it reported no data from independent mortgage companies, and it showed no data on loan rejections. Beginning in 1991, people will be able to see comprehensive data that might indicate loan biases against minorities or lower income people; these data will be available by individual financial institutions. This is good news in itself. Perhaps other federal (and state) agencies will follow HMDA's example.

Hennepin and Ramsey counties have acquired geographic information systems (GIS) that will improved capabilities for address matching and mapping. The two central cities also are developing GIS capability. The Minneapolis Planning Department was quite helpful in the MICAH project and produced maps and property listings that were critical to the success of the project. As more agencies become aware of these GIS capabilities and as the GIS user interface becomes more friendly, there is hope that more data will be forthcoming.

RECOMMENDATIONS FOR IMPROVING MATTERS

There are two basic issues that government should address in creating small area data. The first of these is making a commitment to generate regular reports. The second is installing the technology and commitment to generate additional small area summary reports on request. Both issues require enlightened leadership from elected officials and government managers.

The need for different types of small area data is probably endless, but some types are generic and should be reported-out on a regular basis. Each community may have a different list of what is important to it, but the list will probably look like Table 2. It is important for each community to develop its own list of what items will be reported and with what frequency. Then a lead agency needs to be assigned the task of collecting and disseminating that information.

Of course the biggest weakness of this approach is the lack of current small area data for population, income, and unemployment. Good small area population estimates and projections now are available from private firms and would be a good investment for a city or metropolitan government. In the past, these data were much inferior to estimates based on local knowledge (Russell 1981). Even now the variability among estimates and projections by individual firms is quite large, especially for smaller reporting units (*e.g.* tracts) and rapidly growing areas (ICSC 1986). It behooves purchasers to select a firm with a good track record and one using sound methodologies, such as basing its estimates on housing units and its projections on cohort survival (Finch 1991). Critical to the use of housing unit data is good information on the number of people per household; keeping this information current is one of the purposes of the Ann Arbor Household Survey (Bohl 1990).

Income and unemployment data exist in departments of state government. Pressure should be brought to bear on those departments to meet the needs for small area data. Private firms sell income data, but those products suffer from the lack of good local data upon which to base any rational variation within the city.

Printed reports are useful, but better approaches exist for distributing data on a regular basis. One approach might be hypermedia. St. Louis has developed a

hypermedia system for distributing diverse data about its riverfront (Kindleberger 1989). The system includes a structured browser, a great deal of information, and the ability to view the landscape at different scales. Those who have used GIS Tutor and played with its spatial query examples know how easy it is to use Hypercard to point at a city on a map and get information about it (Green and Raper 1989).

Preconceived reports will never be enough because of changing interests and needs. Government needs to make the commitment to meet all the reasonable requests for *ad hoc* small area data summaries. This means investing in the technology to do address matching and in the information systems that can summarize individual records into small area totals. More important, bureaucratic attitudes need to be changed so that agencies become more willing to consider those requests, and even to take pride in meeting them.

Bureaucracies change only when hit hard, and the best hitters are elected officials--who are supposed to provide leadership and make government responsible to its constituency. An enlightened mayor or council member can work wonders. This general statement about the source of change does not rule out the "white knight," the person inside the bureaucracy who is enlightened and who can change things.

Whatever the source, it is time for government to keep the promises made two decades ago and begin to develop and distribute small area data, so we don't have to wait another ten years to find out what is happening to our cities. The technology is there; the necessary operational data are there. What is lacking are the vision and the will to utilize operational data for additional purposes, to support management and decision making. It's time to keep our promises!

REFERENCES

Barrette, Lynne C., 1989, "Santa Monica Becomes an Electronic City," *Exemplary Systems in Government,* URISA, 177-183.

Bohl, Frederick L., 1990, "The Creation of a Local Demographic Database: A Case Study of the Ann Arbor Household Survey," URISA *Proceedings,* IV, 199-210.

Colley, Donald G., 1974, "Some Considerations in the Development of a Social Change Index to Reflect Current and Comparative Social Change Conditions in Particular Rhode Island Communities," URISA *Proceedings,* II, 406-429.

Cooke, Donald F. and William H. Maxfield, 1967, "The Development of a Geographic Base File and its Uses for Mapping," URISA *Proceedings*, 207-218.

Dangermond, Jack, 1988, "Who is Designing Geographic Information Systems for the Public?" URISA *Proceedings,* III, 37-45.

Finch, Joan G., 1991, Director of Research and Planning, Dayton Hudson Corporation, personal communication, March 27.

Green, Nick and Jonathon Raper, 1989, *GIST: The Geographic Information Systems Tutor,* version 1.03, Birkbeck College, London.

Greer, James L., 1991, "Neighborhood Information Systems," Urban Data Systems, Inc., Red Hook NY.

Heath, Richard L., Blake Graham, and David Nelson, 1986, "Using Geoprocessing to Access Intercensal Small Area Demographic Data," unpublished paper presented at 1986 URISA conference.

ICSC, 1986, "Demographic Data Comparisons," unpublished report produced under the auspices of the International Council of Shopping Centers.

Kindleberger, Charles P., 1989, "Hypermedia Systems: Implications for State Provincial and Local Government," URISA *Proceedings,* III, 142-152.

McMaster, Lorne, 1987, "Edmonton's Neighborhood Fact Sheets Program," URISA *Proceedings,* IV, 103-114.

Mitchell, William, 1971, "Work Session Summary-U.S.A.C. Research Findings," URISA *Proceedings*, 137-138.

Rogers, Clark, 1980, "The Neighborhood Movement-Focus of the 80's" symposium presented at 1980 URISA Conference.

Russell, Cheryl, 1981, "Evaluating Small-Area Estimates and Projections," *American Demographics,* April, 30-32.

Stuart, Kimberly, 1990, "Whittier Neighborhood Housing Profile," Center for Urban and Regional Affairs, University of Minnesota, Minneapolis.

Wolter, Patti, 1990, "Chicago Metropolitan Information Center Offers Tailor-Made Research Services," *The Information Works,* Center for Neighborhood Technology, Chicago, XIII(5) 23.

ACKNOWLEDGEMENTS

The work of two graduate students was very important to collecting and organizing material presented in the paper. Paul McLean did the inventory of existing published and unpublished reports of small area data. Kimberly Stuart worked on the neighborhood housing inventory. Frederick Smith of CURA contributed to both projects. Their assistance was appreciated and is hereby acknowledged.